



OIPE

RAW SEQUENCE LISTING

DATE: 04/03/2002

PATENT APPLICATION: US/10/068,134

TIME: 16:22:36

Input Set : N:\Crf3\RULE60\10068134.raw

Output Set: N:\CRF3\04032002\J068134.raw

1 <110> APPLICANT: Kapeller-Libermann, Rosana
 2 MacBeth, Kyle J.
 3 Williamson, Mark
 4 <120> TITLE OF INVENTION: 22012, A Novel Human Carboxypeptidase
 5 <130> FILE REFERENCE: 5800-38
 6 <140> CURRENT APPLICATION NUMBER: 10/068,134
 7 <141> CURRENT FILING DATE: 2002-02-06
 9 <150> PRIOR APPLICATION NUMBER: US/09/345,469
 10 <151> PRIOR FILING DATE: 1999-06-30
 13 <160> NUMBER OF SEQ ID NOS: 4
 14 <170> SOFTWARE: FastSEQ for Windows Version 3.0
 16 <210> SEQ ID NO: 1
 17 <211> LENGTH: 492
 18 <212> TYPE: PRT
 19 <213> ORGANISM: Homo sapiens
 20 <400> SEQUENCE: 1
 21 Met Gly Ser Ala Pro Trp Ala Pro Val Leu Leu Leu Ala Leu Gly Leu
 22 1 5 10 15
 23 Arg Gly Leu Gln Ala Gly Ala Arg Arg Ala Pro Asp Pro Gly Phe Gln
 24 20 25 30
 25 Glu Arg Phe Phe Gln Gln Arg Leu Asp His Phe Asn Phe Glu Arg Phe
 26 35 40 45
 27 Gly Asn Lys Thr Phe Pro Gln Arg Phe Leu Val Ser Asp Arg Phe Trp
 28 50 55 60
 29 Val Arg Gly Glu Gly Pro Ile Phe Phe Tyr Thr Gly Asn Glu Gly Asp
 30 65 70 75 80
 31 Val Trp Ala Phe Ala Asn Asn Ser Gly Phe Val Ala Glu Leu Ala Ala
 32 85 90 95
 33 Glu Arg Gly Ala Leu Leu Val Phe Ala Glu His Arg Tyr Tyr Gly Lys
 34 100 105 110
 35 Ser Leu Pro Phe Gly Ala Gln Ser Thr Gln Arg Gly His Thr Glu Leu
 36 115 120 125
 37 Leu Thr Val Glu Gln Ala Leu Ala Asp Phe Ala Glu Leu Leu Arg Ala
 38 130 135 140
 39 Leu Arg Arg Asp Leu Gly Ala Gln Asp Ala Pro Ala Ile Ala Phe Gly
 40 145 150 155 160
 41 Gly Ser Tyr Gly Gly Met Leu Ser Ala Tyr Leu Arg Met Lys Tyr Pro
 42 165 170 175
 43 His Leu Val Ala Gly Ala Leu Ala Ala Ser Ala Pro Val Leu Ala Val
 44 180 185 190
 45 Ala Gly Leu Gly Asp Ser Asn Gln Phe Phe Arg Asp Val Thr Ala Asp
 46 195 200 205
 47 Phe Glu Gly Gln Ser Pro Lys Cys Thr Gln Gly Val Arg Glu Ala Phe

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/068,134

DATE: 04/03/2002

TIME: 16:22:36

Input Set : N:\Crif3\RULE60\10068134.raw

Output Set: N:\CRF3\04032002\J068134.raw

```

48          210          215          220
49  Arg Gln Ile Lys Asp Leu Phe Leu Gln Gly Ala Tyr Asp Thr Val Arg
50  225          230          235          240
51  Trp Glu Phe Gly Thr Cys Gln Pro Leu Ser Asp Glu Lys Asp Leu Thr
52          245          250          255
53  Gln Leu Phe Met Phe Ala Arg Asn Ala Phe Thr Val Leu Ala Met Met
54          260          265          270
55  Asp Tyr Pro Tyr Pro Thr Asp Phe Leu Gly Pro Leu Pro Ala Asn Pro
56          275          280          285
57  Val Lys Val Gly Cys Asp Arg Leu Leu Ser Glu Ala Gln Arg Ile Thr
58          290          295          300
59  Gly Leu Arg Ala Leu Ala Gly Leu Val Tyr Asn Ala Ser Gly Ser Glu
60  305          310          315          320
61  His Cys Tyr Asp Ile Tyr Arg Leu Tyr His Ser Cys Ala Asp Pro Thr
62          325          330          335
63  Gly Cys Gly Thr Gly Pro Asp Ala Arg Ala Trp Asp Tyr Gln Ala Cys
64          340          345          350
65  Thr Glu Ile Asn Leu Thr Phe Ala Ser Asn Asn Val Thr Asp Met Phe
66          355          360          365
67  Pro Asp Leu Pro Phe Thr Asp Glu Leu Arg Gln Arg Tyr Cys Leu Asp
68          370          375          380
69  Thr Trp Gly Val Trp Pro Arg Pro Asp Trp Leu Leu Thr Ser Phe Trp
70  385          390          395          400
71  Gly Gly Asp Leu Arg Ala Ala Ser Asn Ile Ile Phe Ser Asn Gly Asn
72          405          410          415
73  Leu Asp Pro Trp Ala Gly Gly Gly Ile Arg Arg Asn Leu Ser Ala Ser
74          420          425          430
75  Val Ile Ala Val Thr Ile Gln Gly Gly Ala His His Leu Asp Leu Arg
76          435          440          445
77  Ala Ser His Pro Glu Asp Pro Ala Ser Val Val Glu Ala Arg Lys Leu
78          450          455          460
79  Glu Ala Thr Ile Ile Gly Glu Trp Val Lys Ala Ala Arg Arg Glu Gln
80  465          470          475          480
81  Gln Pro Ala Leu Arg Gly Gly Pro Arg Leu Ser Leu
82          485          490

```

84 <210> SEQ ID NO: 2

85 <211> LENGTH: 1653

86 <212> TYPE: DNA

87 <213> ORGANISM: Homo sapiens

88 <220> FEATURE:

89 <221> NAME/KEY: CDS

90 <222> LOCATION: (19)...(1494)

91 <400> SEQUENCE: 2

```

92  cgtccggcgga aaggcgac atg ggc tcc gct ccc tgg gcc ccg gtc ctg ctg      51
93          Met Gly Ser Ala Pro Trp Ala Pro Val Leu Leu
94          1          5          10
95  ctg gcg ctc ggg ctg cgc ggc ctc cag gcg qgg gcc cgc agg gcc ccg      99
96  Leu Ala Leu Gly Leu Arg Gly Leu Gln Ala Gly Ala Arg Arg Ala Pro
97          15          20          25

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/068,134

DATE: 04/03/2002

TIME: 16:22:36

Input Set : N:\Crf3\RULE60\10068134.raw

Output Set: N:\CRF3\04032002\J068134.raw

98	gac ccc ggc ttc cag gag cgc ttc ttc cag cag cgt ctg gac cac ttc	147
99	Asp Pro Gly Phe Gln Glu Arg Phe Phe Gln Gln Arg Leu Asp His Phe	
100	30 35 40	
101	aac ttc gag cgc ttc ggc aac aag acc ttc cct cag cgc ttc ctg gtg	195
102	Asn Phe Glu Arg Phe Gly Asn Lys Thr Phe Pro Gln Arg Phe Leu Val	
103	45 50 55	
104	tgg gac agg ttc tgg gtc cgg ggc gag ggg ccc atc ttc ttc tac act	243
105	Ser Asp Arg Phe Trp Val Arg Gly Glu Gly Pro Ile Phe Phe Tyr Thr	
106	60 65 70 75	
107	ggg aac gag ggc gac gtg tgg gcc ttc gcc aac aac tgg ggc ttc gtc	291
108	Gly Asn Glu Gly Asp Val Trp Ala Phe Ala Asn Asn Ser Gly Phe Val	
109	80 85 90	
110	gcg gag ctg gcg gcc gag cgg ggg gct cta ctg gtc ttc gcg gag cac	339
111	Ala Glu Leu Ala Ala Glu Arg Gly Ala Leu Leu Val Phe Ala Glu His	
112	95 100 105	
113	cgc tac tac ggg aag tgg ctg cgg ttc ggt gcg cag tcc acg cag cgc	387
114	Arg Tyr Tyr Gly Lys Ser Leu Pro Phe Gly Ala Gln Ser Thr Gln Arg	
115	110 115 120	
116	ggg cac acg gag ctg ctg acg gtg gag cag gcc ctg gcc gac ttc gca	435
117	Gly His Thr Glu Leu Leu Thr Val Glu Gln Ala Leu Ala Asp Phe Ala	
118	125 130 135	
119	gag ctg ctc cgc gcg cta cga cgc gac ctc ggg gcc cag gat gcc ccc	483
120	Glu Leu Leu Arg Ala Leu Arg Arg Asp Leu Gly Ala Gln Asp Ala Pro	
121	140 145 150 155	
122	gcc atc gcc ttc ggt gga agt tat ggg ggg atg ctc agt gcc tac ctg	531
123	Ala Ile Ala Phe Gly Gly Ser Tyr Gly Gly Met Leu Ser Ala Tyr Leu	
124	160 165 170	
125	agg atg aag tat ccc cac ctg gtg gcg ggg gcg ctg gcg gcc agc gcg	579
126	Arg Met Lys Tyr Pro His Leu Val Ala Gly Ala Leu Ala Ala Ser Ala	
127	175 180 185	
128	ccc gtt cta gct gtg gca ggc ctc ggc gac tcc aac cag ttc ttc cgg	627
129	Pro Val Leu Ala Val Ala Gly Leu Gly Asp Ser Asn Gln Phe Phe Arg	
130	190 195 200	
131	gac gtc acg gcg gac ttt gag ggc cag agt ccc aaa tgc acc cag ggt	675
132	Asp Val Thr Ala Asp Phe Glu Gly Gln Ser Pro Lys Cys Thr Gln Gly	
133	205 210 215	
134	gtg cgg gaa gcg ttc cga cag atc aag gac ttg ttc cta cag gga gcc	723
135	Val Arg Glu Ala Phe Arg Gln Ile Lys Asp Leu Phe Leu Gln Gly Ala	
136	220 225 230 235	
137	tac gac acg gtc cgc tgg gag ttc ggc acc tgc cag cgg ctg tca gac	771
138	Tyr Asp Thr Val Arg Trp Glu Phe Gly Thr Cys Gln Pro Leu Ser Asp	
139	240 245 250	
140	gag aag gac ctg acc cag ctc ttc atg ttc gcc cgg aat gcc ttc acc	819
141	Glu Lys Asp Leu Thr Gln Leu Phe Met Phe Ala Arg Asn Ala Phe Thr	
142	255 260 265	
143	gtg ctg gcc atg atg gac tac ccc tac ccc act gac ttc ctg ggt ccc	867
144	Val Leu Ala Met Met Asp Tyr Pro Tyr Pro Thr Asp Phe Leu Gly Pro	
145	270 275 280	
146	ctc cct gcc aac ccc gtc aag gtg ggc tgt gat cgg ctg ctg agt gag	915

RAW SEQUENCE LISTING

DATE: 04/03/2002

PATENT APPLICATION: US/10/068,134

TIME: 16:22:36

Input Set : N:\Crf3\RULE60\10068134.raw

Output Set: N:\CRF3\04032002\J068134.raw

```

147      Leu Pro Ala Asn Pro Val Lys Val Gly Cys Asp Arg Leu Leu Ser Glu
148          285                      290                      295
149      gcc cag agg atc acg ggg ctg cga gca ctg gca ggg ctg gtc tac aac      963
150      Ala Gln Arg Ile Thr Gly Leu Arg Ala Leu Ala Gly Leu Val Tyr Asn
151          300                      305                      310                      315
152      gcc tgg ggc tcc gag cac tgc tac gac atc tac cgg ctc tac cac agc      1011
153      Ala Ser Gly Ser Glu His Cys Tyr Asp Ile Tyr Arg Leu Tyr His Ser
154          320                      325                      330
155      tgt gct gac ccc act ggc tgc ggc acc ggc ccc gac gcc agg gcc tgg      1059
156      Cys Ala Asp Pro Thr Gly Cys Gly Thr Gly Pro Asp Ala Arg Ala Trp
157          335                      340                      345
158      gac tac cag gcc tgc acc gag atc aac ctg acc ttc gcc agc aac aat      1107
159      Asp Tyr Gln Ala Cys Thr Glu Ile Asn Leu Thr Phe Ala Ser Asn Asn
160          350                      355                      360
161      gtg acc gat atg ttc ccc gac ctg ccc ttc act gac gag ctc cgc cag      1155
162      Val Thr Asp Met Phe Pro Asp Leu Pro Phe Thr Asp Glu Leu Arg Gln
163          365                      370                      375
164      cgg tac tgc ctg gac acc tgg ggc gtg tgg ccc cgg ccc gac tgg ctg      1203
165      Arg Tyr Cys Leu Asp Thr Trp Gly Val Trp Pro Arg Pro Asp Trp Leu
166          380                      385                      390                      395
167      ctg acc agc ttc tgg ggg ggt gat ctc aga gcc gcc agc aac atc atc      1251
168      Leu Thr Ser Phe Trp Gly Gly Asp Leu Arg Ala Ala Ser Asn Ile Ile
169          400                      405                      410
170      ttc tcc aac ggg aac ctg gac ccc tgg gca ggg ggc ggg att cgg agg      1299
171      Phe Ser Asn Gly Asn Leu Asp Pro Trp Ala Gly Gly Gly Ile Arg Arg
172          415                      420                      425
173      aac ctg agt gcc tca gtc atc gcc gtc acc atc cag ggg gga gcg cac      1347
174      Asn Leu Ser Ala Ser Val Ile Ala Val Thr Ile Gln Gly Gly Ala His
175          430                      435                      440
176      cac ctc gac ctc aga gcc tcc cac cca gaa gat cct gct tcc gtg gtt      1395
177      His Leu Asp Leu Arg Ala Ser His Pro Glu Asp Pro Ala Ser Val Val
178          445                      450                      455
179      gag gcg cgg aag ctg gag gcc acc atc atc ggc gag tgg gta aag gca      1443
180      Glu Ala Arg Lys Leu Glu Ala Thr Ile Ile Gly Glu Trp Val Lys Ala
181          460                      465                      470                      475
182      gcc agg cgt gag cag cag cca gct ctg cgt ggg ggg ccc aga ctc agc      1491
183      Ala Arg Arg Glu Gln Gln Pro Ala Leu Arg Gly Gly Pro Arg Leu Ser
184          480                      485                      490
185      ctc tgagcacagg actggagggg tctcaaggct cctcatggag tgggggcttc      1544
186      Leu
187
188      actcaagcag ctggcgggcag aggggaagggg ctgaataaac gcctggaggc ctggcaaaaa      1604
189      aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa      1653
191 <210> SEQ ID NO: 3
192 <211> LENGTH: 10
193 <212> TYPE: PRT
194 <213> ORGANISM: Artificial Sequence
195 <220> FEATURE:
196 <223> OTHER INFORMATION: consensus sequence for the prolyloligopeptidase
197      family from the Prosite database of protein

```

RAW SEQUENCE LISTING

DATE: 04/03/2002

PATENT APPLICATION: US/10/068,134

TIME: 16:22:36

Input Set : N:\Crif3\RULE60\10068134.raw

Output Set: N:\CRF3\04032002\J068134.raw

```

198      patterns
199 <400> SEQUENCE: 3
200      Ile Phe Gly Gly Ser Asn Gly Gly Leu Leu
201      1          5          10
203 <210> SEQ ID NO: 4
204 <211> LENGTH: 251
205 <212> TYPE: PRI
206 <213> ORGANISM: Artificial Sequence
207 <220> FEATURE:
208 <223> OTHER INFORMATION: consensus sequence for the alpha/beta hydrolase
209      family from the Prosite database of protein
210      patterns
211 <400> SEQUENCE: 4
212      Phe Arg Val Ile Ala Leu Asp Leu Arg Gly Phe Gly Glu Ser Ser Arg
213      1          5          10          15
214      Pro Ser Asp Leu Ala Asp Tyr Arg Phe Asp Asp Leu Ala Glu Asp Leu
215      20          25          30
216      Glu Ala Leu Leu Asp Ala Leu Gly Leu Asp Lys Pro Val Ile Leu Val
217      35          40          45
218      Gly His Ser Met Gly Gly Ala Leu Ala Ala Tyr Ala Ala Lys Tyr
219      50          55          60
220      Pro Glu Glu Arg Val Lys Ala Leu Val Leu Val Ser Thr Pro Ala Pro
221      65          70          75          80
222      Ala Gly Leu Ser Ser Arg Leu Phe Pro Arg Leu Gly Asn Leu Glu Gly
223      85          90          95
224      Leu Leu Leu Ala Asn Phe Phe Asn Arg Leu Ser Arg Ser Val Glu Ala
225      100          105          110
226      Leu Leu Gly Arg Ala Leu Lys Gln Phe Phe Leu Leu Gly Arg Pro Phe
227      115          120          125
228      Val Ser Asp Phe Leu Lys Gln Ala Glu Asp Trp Leu Ser Ser Leu Ala
229      130          135          140
230      Arg Pro Gly Glu Thr Asp Gly Gly Asp Gly Leu Leu Gly Tyr Ala Val
231      145          150          155          160
232      Ala Leu Gly Lys Leu Leu Gln Trp Asp Arg Ser Ala Leu Lys Asp Ile
233      165          170          175
234      Lys Val Pro Thr Leu Val Ile Trp Gly Asp Asp Asp Pro Leu Val Pro
235      180          185          190
236      Leu Lys Ala Ser Glu Lys Leu Ser Ala Leu Phe Pro Asn Ala Glu Val
237      195          200          205
238      Val Val Ile Asp Asp Ala Gly His Leu Ala Leu Leu Glu Lys Pro Glu
239      210          215          220
240      Glu Val Ala Glu Leu Ile Lys Phe Leu Ala Leu Ser Thr Asn Asx Ile
241      225          230          235          240
242      Arg Asp Ala Leu Ser Thr Asn Asx Ile Arg Asp
243      245          250

```

VERIFICATION SUMMARY

DATE: 04/03/2002

PATENT APPLICATION: US/10/068,134

TIME: 16:22:37

Input Set : N:\Crf3\RULE60\10068134.raw

Output Set: N:\CRF3\04032002\J068134.raw